BLS WORKING PAPERS

U.S. DEPARTMENT OF LABOR
Bureau of Labor Statistics

OFFICE OF EMPLOYMENT RESEARCH
AND PROGRAM DEVELOPMENT

Principal-Agent Models of CEO Pay-For-Performance Relationships

David S. Kaplan

Working Paper 318
December 1998

I would like to thank Felipe Balmaceda, Fran Blau, Ken Chay, Todd Keister, Kevin Murphy, Rob Olsen, Andy Schneider, Ramsey Shehadeh, Tim Vogelsang, and seminar participants at the Bureau of Labor Statistics, Cornell University, the University of Illinois, and the University of Texas at Dallas. I especially thank John Abowd and Bob Gibbons for their help on earlier drafts. All remaining errors are mine. The opinions expressed in this paper reflect the views of the author, and do not necessarily reflect the policies of the Bureau of Labor Statistics or the views of other staff members.

Principal-Agent Models of CEO Pay-For-Performance Relationships

David S. Kaplan[1]

Bureau of Labor Statistics[2]

October 8, 1998

Abstract

I estimate CEO pay-for-performance schedules for two purposes. First, the predictions of several agency and sorting models are tested. Second, the validity of a common observation/complaint about CEO compensation policies is examined. The principal empirical finding is that CEOs of firms that are prone to high (stock-market) performance volatility receive compensation schedules that lie entirely above the schedules of other CEOs. This shows that the high levels of pay cannot be compensation for bearing more risk. Hazard models show CEOs of high volatility firms also have lower probabilities of turnover.

[2] Bureau of Labor Statistics, 2 Massachusetts Ave., N.E., Room 4945, Washington, DC 20212-0001, kaplan_d@bls.gov. I would like to thank Felipe Balmaceda, Fran Blau, Ken Chay, Todd Keister, Kevin Murphy, Rob Olsen, Andy Schneider, Ramsey Shehadeh, Tim Vogelsang, and seminar participants at the BLS, Cornell University, the University of Illinois, and the University of Texas at Dallas. I especially thank John Abowd and Bob Gibbons for their help on earlier drafts. All remaining errors are mine.

[2] The opinions expressed in this paper reflect the views of the author, and do not reflect the policies of the Bureau of Labor Statistics or the views of other staff members.

1. Introduction

A common complaint about U.S. executive compensation plans is that they are asymmetric in their treatment of performance: that is, executives are rewarded for above average performance and not punished for below average performance. Crystal expresses this view in his 1991 book, where he states

> Unhappily for the U.S. economy, there are too many CEOs who receive especially high pay but whose companies fail to deliver even average performance.

Note that Crystal does not claim that CEOs of firms that perform poorly more than recoup their losses when performance improves. Crystal's claim is that many CEOs receive especially high pay during periods of poor firm performance. I will present evidence in section 3 of the paper showing that CEO compensation data are somewhat consistent with Crystal's cross-sectional observations.

Crystal also states

> ...a CEO who wants to maximize his income would do well not to aim for steady growth and solid return levels, but rather to aim for highly erratic growth.

The main point of this paper is to advance the hypothesis that Crystal got it backwards. Perhaps CEOs do not receive higher levels of compensation by producing highly erratic growth, but firms prone to highly erratic growth offer compensation contracts that provide high pay for all levels of performance. This could occur because of the need to hire better CEOs, or the need to pay an efficiency wage as in Shapiro and Stiglitz (1984) and Akerlof and Yellen (1986).

Before proceeding any further, I would like to illustrate one candidate explanation for Crystal's

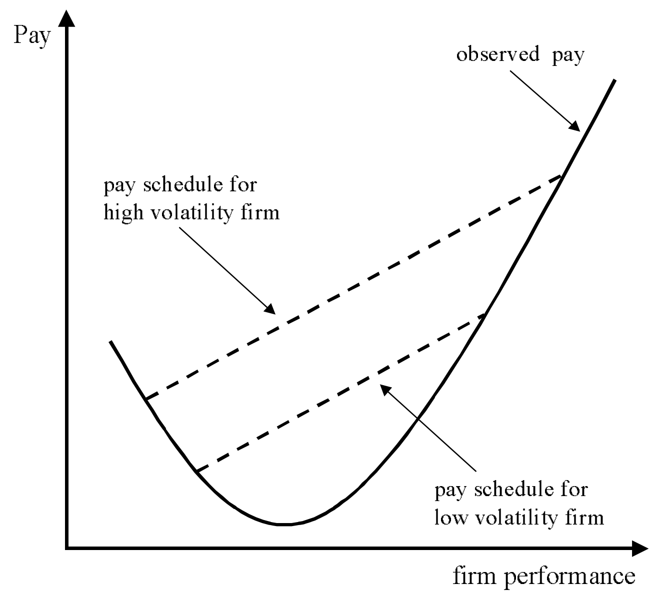

Figure 1: Hypothesized Compensation in a Cross-Section

observations, which is nothing more than a standard omitted variable story. Suppose that all CEO compensation contracts had the same pay-performance slope, but firms prone to high performance variation offered contracts with very high intercepts (base salaries). If we see a firm performing at the bottom of the market-wide performance distribution, we are almost certainly observing a high volatility firm. While, by assumption, this CEO is punished for poor performance as much as any other, we might see high pay relative to **other** CEOs due to the difference in intercepts. In this world, a cross-sectional plot of CEO compensation on realized performance might look like figure 1, in which it appears that CEOs are not punished for low levels of performance. Estimates of the pay-performance slope in cross-sectional data that ignored the differences in intercepts would be biased downward for low levels of performance, and biased upward for high levels of performance. One purpose of this paper is to examine empirically whether Crystal's observations could have been driven by this omitted variable bias. I will also develop agency models that are consistent with this story.

This is a very similar story to the one told in Murphy (1985). Murphy argued that large firms tend to pay their CEOs more, and exhibit lower levels of performance than their smaller counterparts. The omission of a measure of size therefore induces a downward bias in the estimate of the sensitivity of pay to performance. In my story, firms prone to large performance variation offer their executives high pay and account for a disproportionately high percentage of firms with extremely low **and** high performance levels. A negative bias is therefore induced in the estimate of the pay-performance sensitivity for low performance levels, and a positive bias is induced for high performance levels. In section 3, I present empirical evidence on CEO pay providing some support for this conclusion.[1]

I am not the first to argue that firms situated in volatile environments might pay their executives a higher expected level of compensation. Bartlett, Grant, and Miller (1992) and Rose and Shepard (1997) argue that firms prone to high performance variation might be forced to offer their CEOs higher average pay to compensate them for the increased risk these CEOs face. To formalize this argument, consider a principal-agent model in which a risk-neutral principal (board of directors) assigns a risk-averse CEO to a task that yields an output according to the following production function

$$y = k\left(a + \varepsilon\right), \tag{1}$$

where y is the CEO's contribution to firm value, $k > 0$, a is the action (effort) taken by the CEO, and ε has a normal distribution with a mean of zero and a standard deviation of σ.[2] This is essentially the model that drives most studies on executive pay, although the parameter k is not

[1] Aggarwal and Samwick (forthcoming) make a similar point, but they do not mention the upward bias on the pay-performance coefficient for high levels of performance. They also do not explore agency models that would be consistent with the differences in levels of pay that bias the estimated pay-performance slope.

[2] I will often refer to a as effort for simplicity. Higher values of a should really be thought of as actions more consistent with firm value, as opposed to those that bring a private benefit to the CEO.

normally considered. I will further assume a linear wage contract of the form

$$W = S + by, \tag{2}$$

where W is the CEO's wage, S is a base salary and b is the piece rate. Note that, conditional on a, the standard deviation of performance is $k\sigma$, so we have two parameters that determine the "riskiness" of the environment. We will see, however, that these two parameters have important differences in the way they affect the optimal wage contract for the CEO. Suppose further that the CEO's utility function is

$$-\exp\left[-\left(W - \frac{\gamma a^2}{2}\right)\right]. \tag{3}$$

This functional form for utility is used extensively because it yields a closed form solution for expected utility. As I will discuss in section 4, γ is a measure of CEO ability.

First consider the impact of σ, which is assumed to be the source of performance variation in most studies of CEO pay. Note that while σ has no effect on the productivity of the CEO's effort, σ does have an impact on the cost of eliciting effort. This cost arises from the fact that the principal can only elicit high levels of a by linking the CEO's wage to production (y) rather than to a itself. This link between pay and performance can impose enormous risk on a CEO when σ is high. If the principal decides to elicit a high level of a despite a high level of σ, the principal would need to compensate the CEO for bearing high compensation risk. It has been known for a long time that an increase in σ will certainly lead the principal to choose a lower b (inducing the CEO to supply a lower a), but it is possible for S to rise enough so that the expected wage rises.

Assuming $E[W]$ rises, however, we can obtain another empirical prediction. To see this, first note that the standard deviation of the CEO's wage is $bk\sigma$. We already know b will fall in response

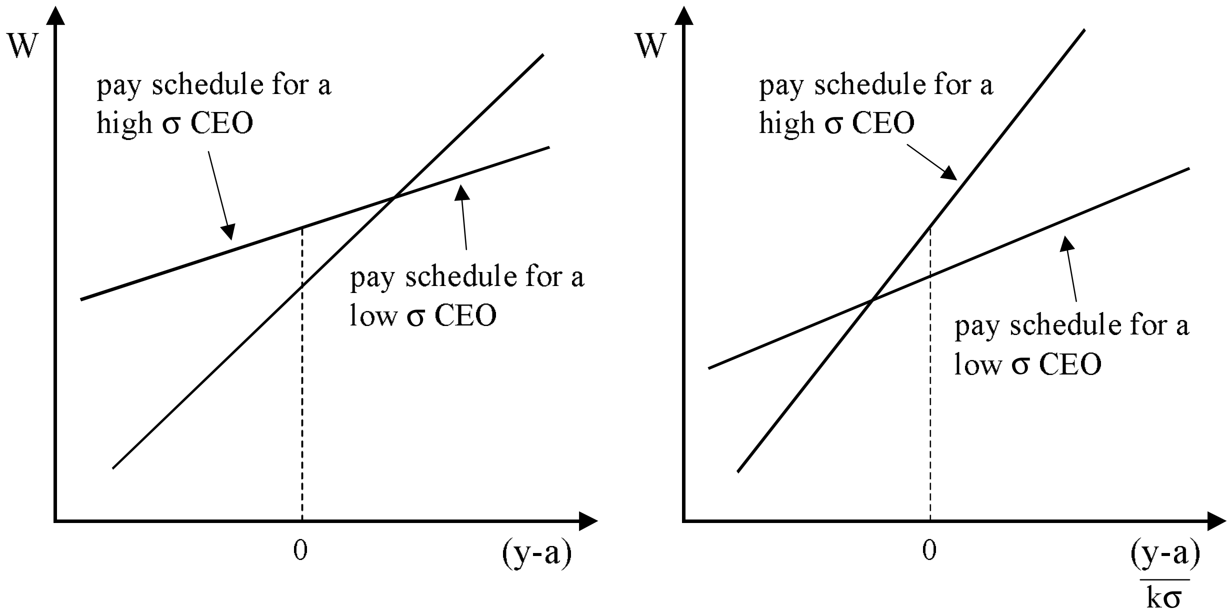

Figure 2: Pay Schedules from the Static Model

to an increase in σ, but suppose b falls so much that the quantity $bk\sigma$ also falls. This would tell us that the CEO's expected wage rises with σ, and the standard deviation of W falls. Since the CEO is also supplying less a, the CEO's expected utility would be increasing in σ, which cannot hold in the optimal contract. Put simply, given an increase in σ, an increase in the expected wage is only possible if the standard deviation of W, $bk\sigma$, also increases. Figure 2 depicts the pay-performance relationships implied by this model if the expected wage is increasing in σ.

Figure 2 also points out the benefits of standardizing the performance measure when plotting compensation schedules. Standardization allows us to evaluate the downside risk associated with the contracts, as well as their compensation levels. When performance is not standardized (as in the left panel), it appears as if the CEO of the high σ firm should be less concerned about the possibility of poor performance, since her compensation is less responsive to performance. This is actually not the case, since low realizations of non-standardized performance are much more likely when σ is high. Looking at the plot with standardized performance (the right panel) shows that the CEO of the high σ firm should be more concerned about ending up in the lower tail of her performance distribution. The importance of standardization will arise again in the empirical section of the paper.

Now consider the effects of the parameter k. It should be clear that k has no impact on the cost of implementing a contract that elicits a fixed level of a. If a principal desires some $\widehat{a}$, it is easy to show that setting $b = \frac{\widehat{a}}{k}$ provides the CEO with the appropriate incentives and same compensation risk, regardless of k. The optimal wage contract does, however, require that the CEO exert higher effort when k is high. This will require that the CEO be compensated both for exerting higher effort and bearing greater compensation risk. Since these pay schedules should look much like those in figure 2, I do not plot them separately.

Neither of these stories does a good job explaining Crystal's observations. Both models predict that the CEOs of high volatility firms should be particularly punished for performance realizations in their lower tail. This is not what Crystal observed. In section 3, I will show that the simple agency stories described above do not accord well with the compensation data used in this paper, in addition to not according well with Crystal's observations.

The outline for the rest of the paper is as follows. Section 2 discusses the data and empirical methodology I use to estimate the effect of performance variability on CEO compensation. I present empirical results in section 3, where I show that CEOs of high volatility firms receive high pay, and exhibit lower probabilities of turnover, than their counterparts in low volatility firms. I also present evidence in section 3 that the failure to account for the effects of volatility on compensation leads to the biases discussed earlier in this introduction. Section 4 shows that a sorting model can only explain my empirical results if volatility is positively correlated with the productivity of CEO actions, in contrast to the standard assumption in the literature. Section 5 presents a dynamic agency model in which pay-for-performance and threat of dismissal are used as incentive aligning tools, and compares their predictions to the results in section 3. Section 5 contains conclusions for the paper.

2. Empirical Implementation

2.1 Objective

In the introduction, I established the reasons I believe it is difficult to reconcile the simplest of agency stories with Crystal's observations. I also demonstrated the importance of analyzing the compensation risk faced by CEOs in assessing the empirical validity of agency models. I address these issues in section 3 by estimating compensation schedules for CEOs, noting how these schedules vary with a measure of the potential for performance variation. This will provide some cross-sectional support for Crystal's observations, and reveal pay-for-performance schedules that are inconsistent with the simple agency stories in the introduction. In sections 4 and 5, I will show that enriched agency models are, in fact, consistent with my empirical results.

2.2 Data

CEO compensation and turnover data come from Standard & Poor's Execucomp covering fiscal years 1992-1996. Unlike most previous studies, Execucomp allows me to include stock-option grants, restricted-stock grants, perquisites, as well as cash compensation in my compensation measure. Returns to common stock were obtained from the Center for Research in Security Prices

(CRSP) and Execucomp. All monetary variables are adjusted using the consumer price index for the last month of the firm's fiscal year. The sample consists of the 3325 CEO-years of data from 821 firms and 1014 CEOs for which the variables could be constructed.[3] Table 1 presents definitions and summary statistics for all of the variables used in the paper. One characteristic that I would like to point out is that age is missing from the data set. In particular, Execucomp only includes the age of those CEOs who are also on the Board of Directors. Appendix A gives more detailed definitions of all variables used in this paper.

2.3 Observed performance variation

An important component of the empirical implementation is my estimate of the potential for performance variation for the firm. As in Garen (1994), my measure is based on firm-by-firm market models, where the estimated standard deviation of the error term is my measure of performance variation. To be more precise, define

$$VWRETD_t = \text{CRSP value weighted market-wide return.}$$

I then estimate

$$ret_{jt} = \beta_{j0} + \beta_{j1} VWRETD_t + u_{jt}, \tag{4}$$

where ret_{jt} is the stock-market return for firm j in year t. I then define

$$sigmm_j = \sqrt{\frac{\sum_t (\varepsilon_{jt})^2}{N_j - 2}}, \tag{5}$$

as my measure of performance variation, where ε_{jt} is the residual from estimating equation (4),

[3] I exclude regulated firms from the sample. See Joskow, Rose, and Shepard (1993) for a discussion of the CEO pay practices of regulated firms.

and N_j is the number of years of return data for firm j.[4]

3. Empirical Results

3.1 Compensation

Recall that the static agency models of the introduction imply that if CEOs in high volatility firms receive higher levels of expected compensation, they should also face greater downside risk. As depicted in figure 2, the pay of the CEO of a high volatility firm should be lower than the pay of a CEO of a lower volatility firm, when both firms perform near the bottom of their respective performance distributions. Figure 2 also shows that we need to standardize the performance measure in order to see this property when looking at compensation schedules.

In order to address this issue, I divided my sample into three (roughly) equal components based on their values for $sigmm$. I then defined dummy variables, L_j, M_j, and H_j to equal one if the firm's value for $sigmm$ is in the lowest, middle or highest $sigmm$ group respectively. Table 2 presents the estimates of the following model:

$$\begin{aligned} Log(\text{Compensation})_{jt} = {} & \sum_{t=1992}^{1996} \beta_{0t} + \beta_1 L_j + \beta_2 M_j + \beta_3 H_j + \beta_4 L_j \frac{(\varepsilon_{jt})}{sigmm_j} + \\ & \beta_5 M_j \frac{(\varepsilon_{jt})}{sigmm_j} + \beta_6 H_j \frac{(\varepsilon_{jt})}{sigmm_j} + \\ & \beta_7 \left(\text{tenure}\right)_{jt} + \beta_8 Log(\text{Assets})_{jt} + \beta_9 Log\left(\text{Sales}\right)_{jt} + \mu_{jt}, \end{aligned} \tag{6}$$

where $Log(\text{Compensation})_{jt}$ is the Log of total compensation awarded to the CEO of firm j in year t, tenure$_{jt}$ is the CEO's tenure at the beginning of the fiscal year, $Log(\text{Assets})_{jt}$ is the Log of the book value of assets, $Log\left(\text{Sales}\right)_{jt}$ is the Log of sales, and ε_{jt} is the residual from estimating equation (4). Under this specification, the intercepts are allowed to differ across groups, as can the

[4] I exclude firms for which I could not obtain at least nine years of stock return data.

coefficients on standardized return.

Assuming performance as predicted from estimating equation (4), a CEO of a medium $sigmm$ firm is predicted to earn 6% more than a CEO of a low $sigmm$ firm, although this difference is significant only at the 0.10 level. The coefficients on standardized return are nearly identical, indicating that this relationship is roughly constant across performance realizations. Again assuming performance as predicted from estimating equation (4), a CEO of a high $sigmm$ firm is predicted to earn 26% more than a CEO in a low $sigmm$ firm and 19% more than a CEO in a medium $sigmm$ firm. Both of these results are significant at the 0.01 level. While the coefficient on standardized return is higher for CEOs of high $sigmm$ firms, the estimated compensation schedules do not cross within 3 standard deviations of predicted performance. Figure 3 displays these results graphically.

It is important to recognize that the estimates in table 2 cannot be interpreted as parameters from the theoretical models. In order to run standard compensation models, *i.e.*, the Log of compensation (rather than simply compensation) as the dependent variable and the rate-of-return as the performance measure, I lose the ability, for instance, to interpret the coefficient on standardized performance as an estimate of $bk\sigma$. We can, however, take a somewhat looser approach. It is clear that CEOs of high $sigmm$ firms receive higher expected compensation than CEOs of low or medium $sigmm$ firms. Either of the static agency models presented earlier can handle this possibility, as long as the CEO of a high $sigmm$ firm does worse than her lower variance counterpart when both exhibit poor performance. The data seem more consistent, however, with the notion that the pay-performance schedules of high $sigmm$ firms are above those for low and medium $sigmm$ firms for all relevant standardized performance outcomes.

Still another potential criticism of the above results is that they are obtained by imposing lin-

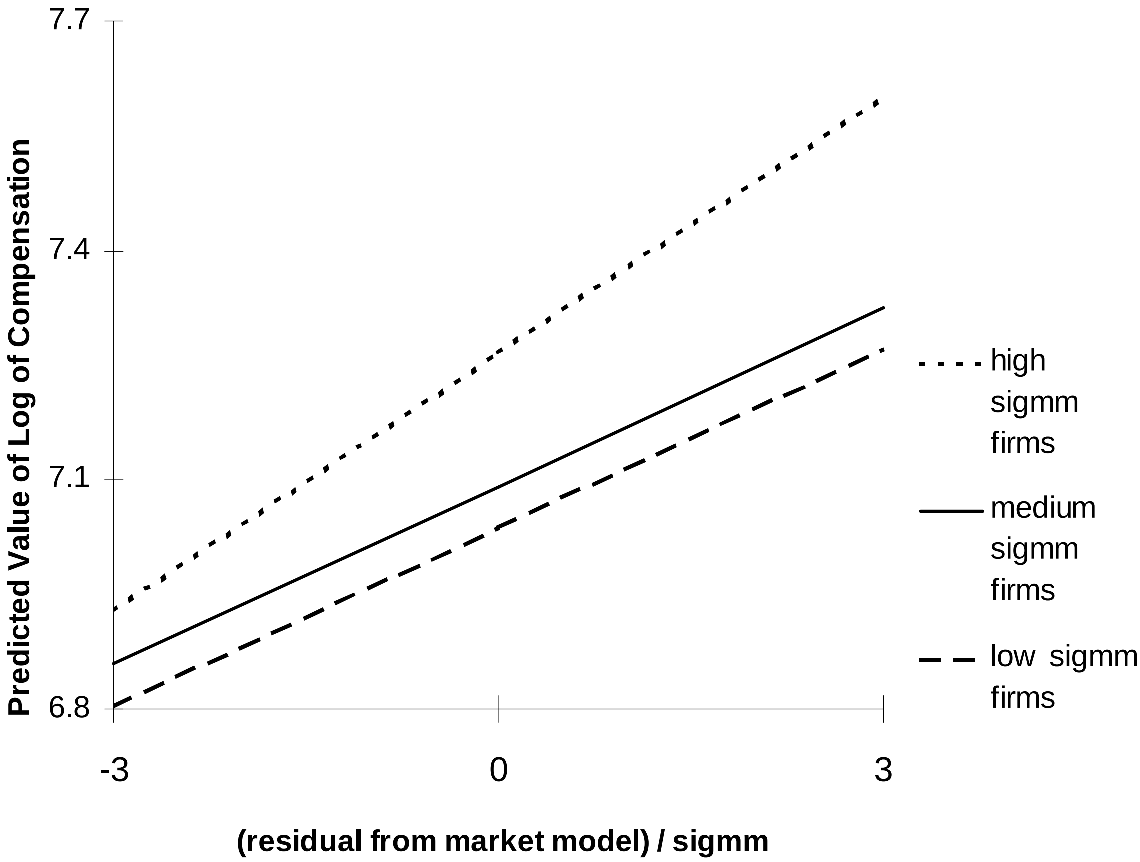

Figure 3: Estimated Pay Schedules

earity on the schedules, which has little theoretical or empirical justification. In order to allow for a non-linear functional form, I assigned each residual from the estimation of each firm's market model to the appropriate quartile of the **firm's** distribution over time. The effect of $sigmm$ on Log(Compensation) was allowed to differ by quartile. Specifically, I estimated

$$\begin{aligned} Log(\text{Compensation})_{jt} = & \sum_{t=1992}^{1996} \beta_{0t} + \sum_{q=2}^{4} \beta_{1q} Q_{jqt} + \sum_{q=1}^{4} \beta_{2q} \left(sigmm\right)_j Q_{jqt} + \\ & \beta_3 \left(\text{tenure}\right)_{jt} + \beta_4 Log(\text{Assets})_{jt} + \beta_5 Log\left(\text{Sales}\right)_{jt} + \mu_{jt}, \end{aligned}$$

where $Q_{jtq} = 1$ if the firm's current residual from the market model is in the q^{th} quartile of the firm's distribution of residuals over time. These results are reported in table 3.

Table 3 shows that a CEO of a high $sigmm$ firm is predicted to receive a higher level of compensation than her counterpart in a low $sigmm$ firm that performed in a similar position in its distribution. In fact, this effect does not dissipate for firms performing in the bottom quartile of their return distributions. Once again, these results indicate that the main difference between the compensation schedule of a CEO of a volatile firm and the schedule of a CEO of a less volatile firm is that the schedule of the first CEO is simply shifted up, providing compensation that is above her counterpart in the less volatile firm for any comparable part of their respective performance distributions. This is inconsistent with the static agency models outlined in the introduction.

Do these results give us any insight into Crystal's observations on CEO pay? My first step in addressing this question is to estimate an analog of figure 1, by estimating a cross-sectional compensation equation that ignores my measure of volatility $(sigmm)$. Turning to table 4, I present estimates of the following model

$$\begin{aligned} Log(\text{Compensation})_{jt} = & \sum_{t=1992}^{1996} \beta_{0t} + \sum_{q=2}^{4} \beta_{1q} MQ_{jqt} + \beta_2 \left(\text{tenure}\right)_{jt} + \\ & \beta_3 Log(\text{Assets})_{jt} + B_4 Log\left(\text{Sales}\right)_{jt} + \mu_{jt}, \end{aligned} \tag{7}$$

where $MQ_{jqt} = 1$ if the firm's current return is in the q^{th} quartile of the return distribution for all firms in the compensation sample for year t.

Note that the predicted difference in the Log of total compensation between CEOs of firms in the lowest and second lowest quartiles of the market-wide distribution of stock returns is positive and significant at the 0.05 level. Crystal's quote notwithstanding, it does seem that CEOs of poorly performing firms receive lower compensation than their colleagues in better performing firms. We will see shortly, however, that accounting for differences in performance volatility will increase the magnitude of this estimate and make it significant at a more stringent level.

Turning now to table 5, we see that high $sigmm$ firms due, in fact, disproportionately represent the upper and lower quartiles of the market-wide distribution, as assumed in figure 1. As discussed earlier, this should lead us to expect that including $sigmm$ in the compensation equation should strengthen my estimate of the pay-performance link at the low end of the performance distribution, and weaken it at the high end. Table 6 shows the results of including $sigmm$ when estimating equation (7). Note that the inclusion of $sigmm$ increased the magnitude of the predicted difference in the Log of Compensation between performance in the second and bottom quartiles of the market distribution for the year (from 0.07 to 0.10) and strengthened the level of significance for this result (from 0.05 to 0.01). Note also that the predicted difference between performance in the highest and third quartiles is 0.10, compared to 0.13 in table 2. Both of these changes from table 4 to table 6 accord well with the explanation of Crystal's observations outlined in the introduction, namely that his failure to account for differences in the intercepts of compensation contracts confounded his observations on their slopes. None of the differences in coefficients between table 4 and table 6, however, are statistically significant.

3.1.1 Compensation Risk From Stock Holdings?

Hall & Liebman (1997) show that the component of pay-for-performance due to CEO stock holdings is much larger than the components due to year-end bonuses and stock-option holdings, particularly over bad performance realizations. I now turn to the question of whether the pay premiums for CEOs of volatile firms merely reflect compensating differentials for the high volatility of their stock holdings. Tables 7 and 8 address this issue.

Recall from table 2 that CEOs of high $sigmm$ firms are predicted to earn much more than their colleagues in low or medium $sigmm$ firms. The predicted difference between medium and low $sigmm$ firms is much smaller and only marginally statistically significant. Tables 7 and 8 therefore focus on CEOs of high $sigmm$ firms. In particular, compensation of CEOs of high $sigmm$ firms that comes in the form of restricted-stock grants is excluded from the compensation measure. This is the only difference between tables 2 and 7, and the only difference between tables 3 and 8. Examination of tables 7 and 8 reveals that they are quite similar to tables 2 and 3, despite the exclusion of stock grants to CEOs of high $sigmm$ firms. CEOs of high $sigmm$ firms are still predicted to earn more than those of low or medium $sigmm$ firms given any reasonable draw from a standardized performance distribution.

While I do not contest the Hall and Liebman claim that CEO stock holdings can lead to enormous volatility in wealth, tables 7 and 8 show that the compensation results presented earlier must be picking up more than compensating differentials for differences in wealth volatility. To see this, suppose CEO A earns \$500,000 per year in salary with no stock grants, while CEO B earns \$1,000,000 per year in salary in addition to extensive stock grants. Assuming equal career lengths (which will be addressed in the next subsection), CEO B's lifetime compensation is guaranteed to be higher than that of CEO A. This is true despite the fact that CEO B, who accumulates more and more stock over her career, can face dramatic reductions in wealth following a bad year.

This analysis does point out one important limitation of my methodology. While the models in the following sections do have specific predictions on the amount of compensation risk imposed on the CEOs, I hesitate to try to accept or reject these implications in the data without a more careful analysis of stock and option holdings that is beyond the scope of this paper. This would be particularly difficult to implement using the Execucomp database, since option holdings are not observable. I will, however, note when a model predicts that CEOs of a high volatility firms receive compensation levels high enough to eliminate compensating differentials as an explanation.

3.2 Turnover

Another potential criticism of my analysis is that CEOs of high volatility firms receive higher pay to compensate them for higher risk of dismissal. I investigate this empirically by estimating a proportional hazard model of CEO tenure duration. Table 9 shows the results of estimating

$$h_j(fy) = \alpha(fy)\exp\left[\begin{array}{l}\sum_{t=1992}^{1996}\beta_{0t} + \beta_1 sigmm_j + \beta_2\frac{(\varepsilon_{jt})}{sigmm_j} + \\ \beta_3 Log(\text{Assets})_{jt} + \beta_4 Log\left(\text{Sales}\right)_{jt}\end{array}\right], \tag{8}$$

where ε_{jt} is the residual from estimating equation (4), $h_j(fy)$ is the turnover hazard for the CEO of firm j with tenure fy, and $\alpha(fy)$ is a baseline hazard. A CEO spell is coded as ending at the end of the last full fiscal year of service. The advantage of assuming this form for the hazard function is that consistent estimates can be obtained for the coefficients of interest without estimating (or assuming a functional form for) $\alpha(fy)$.

Looking at table 9, we see that CEOs of firms with higher values of $sigmm$ have lower probabilities of turnover than their counterparts in lower $sigmm$ firms, although this result is only significant at the 0.05 level. In section 5, I discuss whether this result is consistent with an efficiency-wage model in which pay-for-performance and threat of dismissal are both used to motivate the CEO. Dividing performance into quartiles as in table 4 yielded insignificant results.

4. Sorting of Ability

Perhaps the simplest explanation of the results presented above and Crystal's observations is that the CEOs of firms prone to high volatility are better (more productive) than their colleagues in low volatility firms. In this section, I examine whether or not the simple agency model I outlined in the introduction does indeed predict that high quality CEOs should sort to firms prone to high performance variation.

In this section, I will once again assume production is of the form shown in equation (1), and utility is of the form shown in equation (3). Note that the parameter γ is a good measure of CEO ability. When γ is low, the CEO's marginal cost of effort is also low. Assuming identical incentive contracts, CEOs with low values of γ will always exert more effort. A monotonic transformation of expected utility yields

$$S + bka - \frac{\gamma a^2 + b^2 k^2 \sigma^2}{2},$$

It is easy to show that, assuming the optimal linear contract, $b^* = \frac{1}{1+r\sigma^2}$ and $a^* = \frac{k}{\gamma}$.

In order to determine the efficient sorting of CEOs who differ in γ, I start by calculating the total surplus generated by a CEO-firm match. This surplus, as described in Gibbons (1997), is total output (y) minus disutility of effort $\frac{\gamma a^2}{2}$ minus the cost due to risk aversion $\frac{b^2 k^2 \sigma^2}{2}$, which simplifies to

$$TS = \frac{k^2}{2\gamma\left(\gamma\sigma^2 + 1\right)}.$$

Since

$$\frac{\delta^2\left(TS\right)}{\delta\sigma^2\delta\gamma} = \frac{k^2\sigma^2}{\left(\gamma\sigma^2 + 1\right)^3} > 0,$$

we know that high γ (low ability) CEOs should sort to high σ^2 firms. While high γ CEOs always generate lower total surplus than low γ CEOs $\left(\frac{\delta(TS)}{\delta k} < 0\right)$, this difference is smaller in high σ^2 firms. This does not accord well with the empirical results presented earlier if we assume differences in estimated variance were driven by differences in σ^2. Note however that since

$$\frac{\delta^2\left(TS\right)}{\delta\gamma\delta k} = \frac{-k\left(1 + 2\gamma\sigma^2\right)}{\gamma^2\left(\gamma\sigma^2 + 1\right)^2} < 0,$$

low γ (high-ability) CEOs should sort to high k (high-variance) firms. This does accord well with the empirical results if we assume differences in estimated variance were driven by differences in k.

The intuition for these results is quite simple. Suppose that high performance variation in the data is associated with a more difficult contracting environment, *i.e.*, high σ^2. As σ^2 rises, a firm finds the provision of high-powered incentives more costly. It is therefore inefficient to assign those CEOs who are most responsive to high-powered incentives (low γ) to environments where the provision of these incentives is prohibitively costly (high σ^2). If, however, the performance variation we see in the data is associated with higher productivity of effort, *i.e.*, high k, high-quality CEOs should sort to high volatility firms, simply because high effort is so important.

In summary, if the estimated differences in performance variation correspond to differences in k (productivity of effort) rather than σ^2 (the cost of eliciting effort), than an ability sorting model can be reconciled with the empirical results in section 3. In the former case, firms prone to higher variability choose to hire higher quality CEOs, and therefore pay them higher levels of compensation.

5. The Efficiency-Wage Model

This section will show that when firms prone to high performance variation also have a higher productivity of CEO actions, an efficiency-wage model can be used to explain Crystal's observations. An efficiency-wage model in which differences in performance variation reflect differences in the cost of eliciting effort (which is the typical assumption), does not yield sharp enough predictions to give much guidance to the results in section 3.

5.1 Setup

Once again, utility will be as in equation (3), although I will add to it a positive constant C so that I can normalize the value of the CEO's next best alternative (in utility terms) to zero. The CEO discounts future utility by a factor of $0 < \delta < 1$. Each period, the firm assigns the infinitely lived CEO to a project with a production function as in equation (1). The firm is a risk-neutral profit maximizer, where maximizing the present discounted stream of profits will turn out to be equivalent to maximizing period-by-period profits.

In addition to a linear wage contract as in equation (2), the firm can use the threat of dismissal as a motivational tool. Dismissal will occur when the current performance level drops below some threshold.[5] The firm can replace a dismissed CEO with an identical one without cost. The firm and

[5] The dismissal contract can be an implicit one, in the spirit of Bull (1987).

the CEO play a repeated game with the following timing:[6]

1. The firm sets the compensation contract, which can depend only on the realization of firm performance. The firm also sets a rule by which the CEO can be dismissed at the end of the period, after compensation for the period has taken place.[7]
2. The CEO chooses her level of effort (a).
3. Firm performance is realized. Compensation occurs as specified in the contract and the CEO is dismissed or not as specified by the firing rule.
4. If the CEO is not dismissed, she returns in the next period. If she is dismissed, an identical CEO is hired for the start of the next period. Either way, the game returns to step 1.

I will restrict my search for equilibria to the steady-state equilibrium that maximizes firm profits.[8]

If the CEO believes that the firm will play the same strategy in each period, it is indeed optimal for the firm to play the same strategy every period.[9] The firm's problem can therefore be written as one in which it sets a compensation contract that is constant over time. The firm also chooses a firing rule (which will be derived later) that generates $p(a)$, the probability that the CEO will be fired for a given a. Before going further, it would be helpful to develop some intuition at this point. We might consider two strategies that a firm has to elicit effort. One might be to offer a compensation contract that offers a steep pay-for-performance slope (high b), which provides the CEO with high pay only following good performance. An attractive feature of this contract is that it provides high within-period incentives: that is, the CEO can increase her pay this period by exerting more effort. Another strategy the firm might try is to offer a wage contract with a high base salary (S), and a very low pay-for-performance slope. In this case, the CEO's pay in the current period will not be affected much by her effort. This scheme, however, might motivate the CEO to exert high effort when combined with an appropriate firing rule, since the CEO will work

6 Macleod and Malcomson study a similar model, with endogenous threat of a dissolution of the relationship by both parties. They do not, however, allow for uncertainty, which drives the empirical predictions of this section.

7 It is easy to add additional signals of CEO effort to enter into the firing rule, as long as these signals are also normally distributed and conditionally independent. These additions do not change any of the results.

8 By a steady-state equilibrium, I mean an equilibrium in which the firm and the CEO play strategies that are constant over time.

9 To see this, note the firm is in the same position at the start of each period.

quite hard to preserve her highly-paid position.

Let us now consider the CEO's problem, given some probability of dismissal, $p(a)$, and a compensation contract of the form in equation (2).[10] The problem is

$$\max_{a}\sum_{t=0}^{\infty}\left\{\left[\int_{-\infty}^{\infty}-\exp[-\left(S+bk(a+\varepsilon)-\frac{\gamma a^2}{2}\right)]f(\varepsilon,0,\sigma)d\varepsilon+C\right]\delta^t(1-p(a))^t\right\}. \quad (9)$$

where $f(\varepsilon,\mu,\sigma)$ is the probability density function of a normal variable ε with mean μ and standard deviation σ.[11] Denote $F(\varepsilon,\mu,\sigma)$ as the cumulative distribution function of the same variable ε. The firm will choose to fire the CEO when y is below some value ζ, where ζ is chosen by the firm. This makes the firing probability for a given a and ζ, $p(a;\zeta)=F(\zeta,a,\sigma)=\Phi\left[\frac{\zeta-a}{\sigma}\right]$, and $p'(a;\zeta)=\frac{\partial F(\zeta,a,\sigma)}{\partial a}=\frac{-\phi\left[\frac{\zeta-a}{\sigma}\right]}{\sigma}$ where Φ and ϕ are the cumulative and probability distribution functions of a standard normal variable respectively. The firm will choose the firing rule to maximize

$$\frac{-\delta p'(a;\zeta)}{1-\delta+\delta p(a;\zeta)}=\left(\frac{1}{\sigma}\right)\frac{\delta\phi\left[\frac{\zeta-a}{\sigma}\right]}{1-\delta+\delta\Phi\left[\frac{\zeta-a}{\sigma}\right]}.$$

Equation (9) simply shows the problem of maximizing the discounted level of expected utility for the CEO. In period 0, the CEO's expected utility is

$$\int_{-\infty}^{\infty}-\exp[-\left(S+bk(a+\varepsilon)-\frac{\gamma a^2}{2}\right)]f(\varepsilon,0,\sigma)d\varepsilon+C.$$

In period 1, the CEO will receive the same level of expected utility with probability $(1-p(a))$ and utility of zero with probability $p(a)$. Of course, the CEO discounts this utility by a factor of δ. Continuing on, the CEO receives this same level of expected utility in each period with probability

[10] Since the executive has the same problem at the start of each period (unless she is fired), it is not an imposition to force her to choose the same level of a each period.

[11] The CEO's first-order condition uniquely defines her optimal level of a.

$(1-p(a))^t$, which she discounts by a factor of δ^t. This problem simplifies to

$$\max_a \frac{-\exp\left[-\left(S+bka-\frac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right]+C}{1-\delta+\delta p(a)}.$$

Note that we have already analyzed a special case of this model in the introduction, since this model collapses down to the static model when $\delta = 0$, *i.e.*, the CEO does not care about the future at all. In this section, I will focus on the opposite extreme, namely the case in which the CEO is extremely patient (δ approaches 1). In particular, I will focus on the empirical implications of this model, noting how they might relate to Crystal's observations and how they relate to the results in section 3. Most of the formal treatment of this model, however, is relegated to appendix B.

Recall that the production function used in this paper allows for two possible explanations for these differences in observed volatility in the data. Firms might differ in their values for σ, which is the typical assumption made in CEO studies like Garen (1994) and Aggarwal and Samwick (1998). As discussed earlier, this would imply that firms differ in their costs of eliciting effort. Firms could also, however, differ in their values of k, which would imply that firms differ in their valuation of high CEO effort. Recall that in the static case ($\delta = 0$), neither of these explanations fit Crystal's observations. We saw in section 3 that they also cannot be reconciled with the compensation data used in this paper. I now turn to the results from the dynamic version of the model.

The relevant comparative statics as $\delta \to 1$, relating to σ are as follows:

$$\frac{\partial a^*}{\partial \sigma} < 0, \frac{\partial b^*}{\partial \sigma} < 0, \frac{\partial (b^* k\sigma)}{\partial \sigma} < 0, \text{ and } \frac{\partial (S^* + b^* ka^*)}{\partial \sigma} \gtrless 0.$$

Unfortunately, this model does not give us a clear prediction on compensation levels, which makes it difficult to test in my compensation data. Let me, however, try to give some intuition for the results. When σ is high, the use of pay-for-performance becomes a particularly unattractive tool, even compared with the threat of dismissal. This is why the piece rate, and even total compensation risk decreases. The sign of the change in expected compensation, given an increase in σ, depends on k (productivity). When k is small, the firm responds to an increase in σ by decreasing a (effort) quite a bit, which allows the firm to pay a lower level of compensation. When k is high, however, the firm resists the temptation to reduce a, since effort is so productive. Since the effectiveness of compensation levels as an incentive aligning tool is decreasing in σ, the firm will need to increase compensation.

The relevant comparative statics as $\delta \to 1$, relating to k are as follows:

$$\frac{\partial a^*}{\partial k} > 0, \frac{\partial (b^*)}{\partial k} < 0, \frac{\partial (b^* k \sigma)}{\partial k} < 0, \text{ and } \frac{\partial (S^* + b^* k a^*)}{\partial k} > 0.$$

The first promising feature we see about these comparative statistics is that expected compensation is predicted to be higher for CEOs in high variability firms. Further, since their compensation risk is predicted to be smaller, it might not be surprising to see that CEOs of poorly performing firms receive high pay, even following very poor performance. The main effect in this model is that, as effort becomes more important, the firm chooses to elicit a higher level of effort through higher pay, which is a standard efficiency-wage argument. This higher level of pay is an incentive tool, not merely compensation for increased effort. It is not hard to show that the utility cost of imposing compensation risk is increasing in expected compensation, which explains the results on compensation variance and the piece rate (b^*).

Before discussing the turnover implications of these models, I would like to discuss their relevance to another study. Kole and Lehn (1996) look at the effect of airline deregulation on CEO pay and turnover. While regulated firms are not in my compensation sample, they do have very low levels of $sigmm$, since their stock returns are quite stable. Kole and Lehn find that, even for CEOs in place long before deregulation, CEO pay rises dramatically after deregulation. The model in this section offers the interpretation that, despite the fact that CEO ability had not dramatically changed, these firms offered higher pay. In particular, these high pay levels could have been used as motivational tools when combined with the threat of dismissal. The firms chose to adopt this policy following deregulation, when it became particularly important to elicit high effort.[12]

5.2 Turnover

The model in this section predicts that the probability of dismissal should not change as either k or σ changes, which is a result entirely due to functional form. Kaplan (1998) presents a model very similar to the one presented here in which the probability of dismissal declines with the parameter that affects the productivity of CEO actions, rather than the cost. The important insight these models yield is that we need not see a higher incidence of dismissal in firms that are relying more heavily on efficiency wages. In the models presented in this section, CEOs of high k (high productivity) firms are held to more stringent dismissal standards that their low k counterparts. CEOs in high k firms do, however, work harder than those in low k firms, equalizing the probability of dismissal.

6. Conclusions

This paper shows that CEOs of firms prone to high performance variation receive higher compensation than CEOs of less volatile firms. Estimated pay schedules show that these differences in pay do not reflect compensating differentials for excessive compensation risk. Hazard models of tenure

[12] Joskow, Rose, and Shepard (1993) and Joskow, Rose, and Wolfram (1996) offer an explanation for low pay in regulated firms based on political pressure to keep salaries down.

duration show they do not reflect compensating differentials for risk of dismissal. I show that an ability sorting model cannot explain these results under the standard assumption that performance variability increases the cost of eliciting effort, without affecting productivity. The results can only be explained by a sorting model if higher performance variation signals higher productivity of CEO actions. Assuming this latter functional form for production, an efficiency-wage model can also explain the compensation results.

The paper may shed light on a common complaint about CEO pay practices. Since firms prone to high variability disproportionately comprise the lower tail of the market-wide performance distribution, and these firms offer their CEOs compensation contracts with high intercepts, it can appear in a cross-section that CEOs are not punished for poor performance. Compensation data lend some support for this explanation.

Appendix A. Data Definitions

$VWRETD_t$.

value weighted average market-wide return (including all distributions). VWRETD is the variable name used by CRSP.

ret_{jt}.

the total shareholder return (including dividends) for the firm's common stock for the fiscal year.

ε_{jt}

the residual for firm j in fiscal year t obtained from estimating equation (4).

$sigmm_j$.

defined from equations (4) and (5).

$sret_{jt}$

scaled shareholder return = $\frac{\varepsilon_{jt}}{sigmm_j}$

$Log(\text{Compensation})_{jt}$.

the Log of total compensation. Total compensation includes salary, bonus, perquisites, payments to cover CEO's taxes, preferential earnings payable but deferred at CEO's election, preferential discounts on stock purchases, the value of restricted stock granted (as reported by the company), the Black-Scholes value of stock-options/SARs granted, long-term incentive payouts, company contributions to benefit plans, split-dollar insurance payments, payments related to termination, change-in-control or severance, and other miscellaneous compensation.

MQ_{jqt}

dummy variable that equals one if firm j is in the q^{th} quartile of the distribution of returns in year t for all firms used in the compensation analysis in year t.

tenure_{jt}.

the CEO's tenure (as CEO, rather than at the firm in any capacity) at the beginning of fiscal year t for firm j.

$Log(\text{Assets})_{jt}$

the Log of total current assets (same as Compustat annual data item 4).

$Log\ (\text{Sales})_{jt}$

the Log of net sales (same as Compustat annual data item 12).

L_j, M_j, and H_j

All firms used in the compensation sample are ranked by their value for $sigmm_j$. L_j is a dummy variable that equals one if firm j is in the lowest third of this distribution. M_j is a dummy variable that equals one if firm j is in the middle third of this distribution. H_j is a dummy variable that equals one if firm j is in the highest third of this distribution.

Q_{jqt}

Each firm has residuals from the estimation of equation (4). Q_{jqt} is a dummy variable that equals one if the residual from year t is in the q^{th} quartile of residuals for firm j.

All monetary variables are adjusted for inflation by using the CPI (urban unadjusted), and are in 1982-84 dollars.

Appendix B. Proof of Compensation Results From Section 5

The purpose of the appendix is to prove the results given in section.5. The CEO's objective is to choose a to maximize the present-discounted expected utility

$$\frac{\int\limits_{-\infty}^{\infty} -\exp[-\left(S+bk(a+\varepsilon)-\frac{\gamma a^2}{2}\right)]f(\varepsilon,0,\sigma)d\varepsilon + C}{1-\delta+\delta p(a)},$$

yielding the first-order condition

$$\begin{aligned}&(b-\gamma a)\exp\left[-\left(S+bka-\tfrac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right]-\\&\tfrac{\delta p'(a)}{(1-\delta+\delta p(a))}\left\{-\exp\left[-\left(S+bka-\tfrac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right]+C\right\}=0.\end{aligned}$$

where $f(\varepsilon,\mu,\sigma)$ is the probability density function of a normal variable ε with mean μ and standard deviation σ. Denote $F(\varepsilon,\mu,\sigma)$ as the cumulative distribution function of the same variable ε. The firm will choose to fire the CEO when y is below some value ζ, where ζ is chosen by the firm. This makes the firing probability for a given a and ζ, $p(a;\zeta) = F(\zeta,a,\sigma) = \Phi\left[\frac{\zeta-a}{\sigma}\right]$, and $p'(a;\zeta) = \frac{\partial F(\zeta,a,\sigma)}{\partial a} = \frac{-\phi\left[\frac{\zeta-a}{\sigma}\right]}{\sigma}$ where Φ and ϕ are the cumulative and probability distribution functions of a standard normal variable respectively. As in section 2, the firm will maximize $\frac{-\delta p'(a;\zeta)}{1-\delta+\delta p(a;\zeta)} = \left(\frac{1}{\sigma}\right)\frac{\delta\phi\left[\frac{\zeta-a}{\sigma}\right]}{1-\delta+\delta\Phi\left[\frac{\zeta-a}{\sigma}\right]}$, which tells us $\zeta < a$ (the firm will only fire the CEO if performance is below the expected level).[13]

Now denote η as the maximum value of $\frac{\delta\phi(x)}{1-\delta+\delta\Phi(x)}$. Given some level of a that the firm wishes to elicit, $\frac{-\delta p'(a;\zeta)}{1-\delta+\delta p(a;\zeta)} = \frac{\eta}{\sigma}$ at the optimal firing rule for the firm.

It is not hard to show that $\lim\limits_{\delta\longrightarrow 1}\eta = \infty$.[14] As $\eta\rightarrow\infty$, the firm can get arbitrarily close to the first-best solution of imposing no risk on the CEO, and merely compensating her for her effort. Changing σor k affects nothing. The question of interest for this appendix is the direction from

[13] The CEO's first-order condition uniquely defines her optimal level of a.

[14] This relies on the same property of the normal density function that was used in Mirrlees (1974).

which the empirically relevant comparative statics approach zero.

Simple calculations show that the firm can only be minimizing cost if

$$ka\frac{\partial a}{\partial S}=\frac{\partial a}{\partial b},$$

which implies that

$$\eta b=\frac{1}{k\sigma}+b\left(bk-\gamma a\right)\sigma \tag{B-1}$$

and

$$\eta^2 b=\frac{\eta}{k\sigma}+\left[\frac{1}{k\sigma}+b\left(bk-\gamma a\right)\sigma\right]\left(bk-\gamma a\right)\sigma. \tag{B-2}$$

These equations tell us that

$$\lim_{\eta\to\infty}\eta\frac{\partial b}{\partial\sigma}=-\frac{1}{k\sigma^2}$$

and

$$\lim_{\eta\to\infty}\eta^2\frac{\partial\left(bk\sigma\right)}{\partial\sigma}=-\gamma a.$$

In order to obtain the results on expected compensation, we need to write the Lagrangian corresponding to minimizing expected cost subject to eliciting some a as follows:[15]

$$L_2=-S-bka+\lambda\left\{\begin{array}{l}\left(b-\gamma a\right)\exp\left[-\left(S+bka-\frac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right]+\\ \frac{\eta}{\sigma}\left\{\exp\left[-\left(S+bka-\frac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right]+C\right\}\end{array}\right\}, \tag{B-3}$$

with first-order conditions:

$$\lambda:\quad\left[\begin{array}{l}\left(b-\gamma a\right)\exp\left[-\left(S+bka-\frac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right]+\\ \frac{\eta}{\sigma}\exp\left[-\left(S+bka-\frac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right]\end{array}\right]=0 \tag{B-4}$$

[15] For high enough η, the constraint that expected utility must be at least zero can be ignored.

$$S: \quad \left[\begin{array}{l} -1+\lambda \text{ multiplied by} \\ \left(\begin{array}{l} -(b-\gamma a)\exp\left[-\left(S+bka-\frac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right]+ \\ \frac{\eta}{\sigma}\exp\left[-\left(S+bka-\frac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right] \end{array} \right) \end{array} \right] = 0 \tag{B-5}$$

$$b: \quad \left[\begin{array}{l} -ka+\lambda \text{ multiplied by} \\ \left(\begin{array}{l} -(b-\gamma a)\left(ka-bk^2\sigma^2\right)\exp\left[-\left(S+bka-\frac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right]+ \\ k\exp\left[-\left(S+bka-\frac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right]+ \\ \frac{\eta}{\sigma}\left(a-bk^2\sigma^2\right)\exp\left[-\left(S+bka-\frac{\gamma a^2+b^2k^2\sigma^2}{2}\right)\right] \end{array} \right) = 0. \end{array} \right] \tag{B-6}$$

Straightforward calculations show that

$$\lim_{\eta\to\infty}\eta^2\frac{\partial C^*}{\partial\sigma} = \eta^2\frac{\partial C}{\partial\sigma} + \left(\eta\frac{\partial C}{\partial a}\right)\left(\eta\frac{\partial a^*}{\partial\sigma}\right) = \gamma^2 a\left(\gamma a-1\right)k\sigma^3.$$

Since the $k=\gamma\alpha$ is a necessary condition for profit maximization, we know

$$\lim_{\eta\to\infty}\eta^2\frac{\partial C^*}{\partial\sigma} > 0 \text{ if and only if } k>1.$$

I now turn to comparative statics relating to changing the parameter k. Note from equation B-2 that

$$\lim_{\eta\to\infty}\eta^2\frac{\partial\left(bk\sigma\right)}{\partial k} = 0$$

and

$$\lim_{\eta\to\infty}\eta^2\frac{\partial\left(bk\sigma\right)}{\partial a} = -\gamma a,$$

which tells us

$$\lim_{\eta\to\infty}\eta^2\frac{\partial\left(b^*k\sigma\right)}{\partial k} = \lim_{\eta\to\infty}\eta^2\frac{\partial\left(bk\sigma\right)}{\partial k} + \lim_{\eta\to\infty}\eta^2\frac{\partial\left(bk\sigma\right)}{\partial a}\left(\frac{\partial a^*}{\partial k}\right) = -\frac{k}{\gamma},$$

which also tells us

$$\lim_{\eta\to\infty}\eta^2\frac{\partial b^*}{\partial k} < 0.$$

The results on expected compensation come much more easily. Simple calculations show

$$\frac{\partial\left[S^{*}+b^{*}ka^{*}\right]}{\partial k}=\frac{k}{\gamma}.$$

As in section 2, it is easy to obtain a turnover prediction from this model. Note that the maximum value of η does not depend on σ. For any σ and a, the cost-minimizing choice for ζ will be such that $\frac{\zeta-a}{\sigma}$ equals some constant. This result does not depend on δ being large, and tells us that the probability of turnover is independent of σ in this model.

References

[1] Akerlof, G.A. and Yellen, J.L. eds. *Efficiency Wage Models of the Labor Market*. New York: Cambridge University Press, (1986).

[2] Aggarwal, R. and Samwick, A.A. "The Other Side of the Tradeoff: The Impact of Risk on Executive Compensation." *Journal of Political Economy* forthcoming.

[3] Bartlett, R.L., Grant J.H., and Miller, T.I. "The Earnings of Top Executives: Compensating Differentials for Risky Business." *Quarterly Review of Economics and Finance*, Vol. 32, No. 1 (Spring 1992): 38-49.

[4] Bull, C. "The Existence of Self-Enforcing Implicit Contracts." *Quarterly Journal of Economics*, Vol. 102, No. 1 (February 1987): 147-159.

[5] Crystal, G.S. *In Search Of Excess: The Overcompensation of American Executives*. New York: WW Norton, 1991.

[6] Garen, J.E. "Executive Compensation and Principal-Agent Theory." *Journal of Political Economy*, Vol. 102, No. 6 (December 1994): 1175-99.

[7] Gibbons, R.S. "Incentives and Careers in Organizations." *In Advances in Economic Theory, Seventh World Congress* Vol. 2, edited by D. Kreps and K, Wallis. Cambridge University Press (1997).

[8] Hall, B.J. and Liebman, J.B. "Are CEOs Really Paid Like Bureaucrats?" mimeo, Harvard University, (March 1997).

[9] Haubrich, J.G., "Risk-Aversion, Performance Pay and the Principal-Agent Problem." *Journal of Political Economy*, Vol. 102, No. 2 (April 1994): 258-76.

[10] Holmstrom, B. "Moral Hazard and Observability." *Bell Journal of Economics*, Vol. 10, No.1 (Spring 1979): 74-91.

[11] Holmstrom, B. and Milgrom, P. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica*, Vol. 55, No. 2 (March 1987): 303-28.

[12] Kaplan, D.S. "Essays on Incentives and Compensation: Theory and Evidence." Ph.D. dissertation, Cornell University (1998)

[13] Joskow, P.L., Rose, N.L., and Shepard, A. "Regulatory Constraints on CEO Compensation." *Brookings Papers on Economic Activity: Microeconomics*, (1993): 1-58.

[14] Joskow, P.L., Rose, N.L., and Wolfram, C. "Political Constraints on Executive Compensation:

Evidence from the Electric Utility Industry." *Rand Journal of Economics*, Vol. 27, No. 1 (Spring 1996): 165-82.

[15] Kole, S and Lehn, K. "Deregulation and the Adaptation of Governance Structure: The Case of the U.S. Airline Industry."mimeo, Simon Graduate School of Business Administration, University of Rochester (1996).

[16] Lazear, E.P. and Rosen, S. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, Vol. 89, No. 5 (October 1981): 841-64.

[17] MacLeod, W.B. and Malcomson, J.M. "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica*, Vol. 57, No.2 (March 1989): 447-80.

[18] Mirrlees, J. "Notes on Welfare Economics, Information, and Uncertainty." in M. Balch, D. McFadden, and S. Wu, eds., *Essays on Economic Behavior Under Uncertainty*. Amsterdam: North Holland Publishing Co., (1974).

[19] Murdoch, E.J. "Executive Compensation and Firm Performance: The Relationship Between Monitoring and the Use of Incentive Contracts." Ph.D. dissertation, University of California, Los Angeles, (1991).

[20] Murphy, K.J. "Corporate Performance and Managerial Remuneration: An Empirical Analysis." *Journal of Accounting and Economics,* Vol. 7, No. 1-3 (April 1985): 11-42.

[21] Rose, R.L. and Shepard, A. "Firm Diversification and CEO Compensation: Managerial Ability or Executive Entrenchment?" *Rand Journal of Economics*, Vol. 28, No. 3 (Autumn 1997): 489-514.

[22] Shapiro, C. and Stiglitz, J.E. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review*, Vol. 74, No. 3 (June 1984): 433-44.

[23] *Standard and Poor's Compustat Execucomp*. McGraw-Hill, Inc. (October 1997).

[24] Stiglitz, J.E. "Incentives, Risk and Information: Notes Towards a Theory of Hierarchy." *Bell Journal of Economics and Management Science*, Vol. 6, No. 2 (Autumn 1975): 552-79.

Table 1: Descriptive Statistics for Execucomp Data (1992-1996)

Variable	Mean	Std. Dev.
sigmm = Estimated standard deviation of the error term from the firm's market model	0.384	0.250
tenure = years as CEO prior to current year	8.460	8.140
ret = stock market rate of return	0.157	0.387
sret = (residual from market model)/sigmm	-0.101	0.846
Log(Sales) = Log of Sales in millions	6.819	1.554
Log(Assets) = Log of Current Assets in millions	7.051	1.752
Log(Compensation)=Log of Total CEO Compensation in thousands	6.971	0.910
Observations	3,325	
CEOs	1,014	
Firms	821	

Table 2: Cross-Sectional Estimate of Standardized Log-Linear Pay Schedules for Low, Middle, and High Variability Firms

Variable	Estimate	Standard Error
medium variability firm	0.055	0.028
high variability firm	0.228	0.034
sret for low variability firms	0.078	0.026
sret for medium variability firms	0.078	0.024
sret for high variability firms	0.112	0.027
Tenure	0.004	0.002
Log(Sales)	0.184	0.016
Log(Assets)	0.211	0.013
year effects included, but not shown		
Observations	3,325	
CEOs	1,014	
Firms	821	
R^2	0.444	

Standard Errors are Calculated Under the Assumption of Arbitrary Heteroscedasticity

Table 3: Effect of Performance Variability on Compensation Across Standardized Quartiles of the Performance Distribution

Variable	Estimate	Standard Error
Performance in 2nd quartile of firm distribution of returns	0.115	0.067
Performance in 3rd quartile of firm distribution of returns	0.097	0.069
Performance in highest quartile of firm distribution of returns	0.200	0.071
sigmm in lowest quartile	0.317	0.126
sigmm in 2nd quartile	0.271	0.100
sigmm in 3rd quartile	0.515	0.119
sigmm in highest quartile	0.397	0.119
Tenure	0.004	0.002
Log(Sales)	0.191	0.015
Log(Assets)	0.204	0.013
year effects included, but not shown		
Observations	3,325	
CEOs	1,014	
Firms	821	
R^2	0.444	

Standard Errors are Calculated Under the Assumption of Arbitrary Heteroscedasticity

Table 4: Cross-Sectional Estimate of CEO Pay-Performance Relationships (ignoring performance variability, non-standardized performance)

Variable	Estimate	Standard Error
Performance in 2nd quartile of market distribution of returns	0.068	0.033
Performance in 3rd quartile of market distribution of returns	0.167	0.034
Performance in highest quartile of market distribution of returns	0.299	0.036
Tenure	0.004	0.002
Log(Sales)	0.179	0.015
Log(Assets)	0.189	0.013
year effects included, but not shown		
Observations	3,325	
CEOs	1,014	
Firms	821	
R^2	0.444	

Standard Errors are Calculated Under the Assumption of Arbitrary Heteroscedasticity

Table 5: Correlations Between Estimated Performance Variability and Quartile of Market Distribution

	sigmm
Performance in lowest quartile of market distribution of returns	0.162
Performance in 2nd quartile of market distribution of returns	-0.149
Performance in 3rd quartile of market distribution of returns	-0.155
Performance in highest quartile of market distribution of returns	0.142

All correlations are significant at the 0.01 level.

Table 6: Cross-Sectional Estimate of CEO Pay-Performance Relationships (accounting for performance variability, non-standardized performance)

Variable	Estimate	Standard Error
Performance in 2nd quartile of market distribution of returns	0.101	0.033
Performance in 3rd quartile of market distribution of returns	0.198	0.034
Performance in highest quartile of market distribution of returns	0.295	0.035
sigmm	0.371	0.064
Tenure	0.004	0.002
Log(Sales)	0.192	0.015
Log(Assets)	0.199	0.013
year effects included, but not shown		
Observations	3,325	
CEOs	1,014	
Firms	821	
R^2	0.452	

Standard Errors are Calculated Under the Assumption of Arbitrary Heteroscedasticity

Table 7: Cross-Sectional Estimate of Standardized Log-Linear Pay Schedules for Low, Middle, and High Variability Firms (Restricted-Stock Grants Excluded for CEOs of High Variability Firms)

Variable	Estimate	Standard Error
medium variability firm	0.051	0.028
high variability firm	0.183	0.034
sret for low variability firms	0.079	0.026
sret for medium variability firms	0.078	0.024
sret for high variability firms	0.104	0.026
Tenure	0.004	0.002
Log(Sales)	0.182	0.015
Log(Assets)	0.209	0.013
year effects included, but not shown		
Observations	3,325	
CEOs	1,014	
Firms	821	
R^2	0.446	

Standard Errors are Calculated Under the Assumption of Arbitrary Heteroscedasticity

Table 8: Effect of Performance Variability on Compensation Across
Standardized Quartiles of the Performance Distribution
(Restricted-Stock Grants
Excluded for CEOs of High Variability Firms)

Variable	**Estimate**	**Standard Error**
Performance in 2^{nd} quartile of firm distribution of returns	0.102	0.066
Performance in 3^{rd} quartile of firm distribution of returns	0.086	0.068
Performance in highest quartile of firm distribution of returns	0.196	0.070
sigmm in lowest quartile	0.256	0.122
sigmm in 2^{nd} quartile	0.239	0.099
sigmm in 3^{rd} quartile	0.482	0.116
sigmm in highest quartile	0.334	0.115
Tenure	0.004	0.002
Log(Sales)	0.188	0.015
Log(Assets)	0.204	0.013
year effects included, but not shown		
Observations	3,325	
CEOs	1,014	
Firms	821	
R^2	0.449	

Standard Errors are Calculated Under the Assumption of Arbitrary Heteroscedasticity

Table 9: Hazard Model of CEO Tenure Duration

Variable	Estimate	Standard Error	Risk Ratio
sigmm	-0.729	0.360	0.482
sret	-0.206	0.084	0.814
Log(Sales)	0.221	0.084	1.247
Log(Assets)	-0.147	0.076	0.863
year effects included, but not shown			
CEO Spells	1,014		
-2(Log Liklihood)	2,030.45		